JOURNAL OF MULTIVARIATE ANALYSIS 7, 525–534 (1977)

A Property of Random Processes with Unit Multiplicity*

ANTHONY EPHREMIDES

University of Maryland

Communicated by T. Hida

The class of random processes with spectral multiplicity equal to one is shown to be dense in various broader classes of processes with finite energy under the norm of $L_2(\Omega \times T; dp \times dt)$. In some special cases constructive procedures for establishing the denseness are outlined. The application of this result to communication systems is then demonstrated.

1. INTRODUCTION

The role of the value of the multiplicity M of a random process in the structural and statistical properties of a random process has been demonstrated by a number of authors [1–7] after the original derivation of the canonical representation by Cramer [8, 9] and Hida [10]. In earlier studies some of the authors [4, 6] have extended the representation to processes with generalized parameter sets, and more recently [11–13] the connection and applicability of these results to engineering problems has been recognized. A comprehensive collection of the major results on multiplicity theory can be found in [14].

In this paper the class of processes under consideration consists of real-valued, scalar random processes with the real line or a closed finite interval $[0, T]$ as parameter set. Let $x(t)$, or just x, be such a process. It is required that $E[x(t)]^2 < \infty$, for all t. Let $H(x, t)$ designate the closed, linear manifold spanned by $x(s)$, $s \leqslant t$. It is further required that $H(x, t)$ be a separable Hilbert space in order for the multiplicity representation to hold; furthermore, we assume that $x(t)$ is purely nondeterministic; a sufficient condition for this is the existence of the left and right mean-square limits $x(t \pm 0)$ for every t and the requirement that $\bigcap_t H(x, t) = \{1\}$. Thus let C denote the class of finite power processes with separable linear spans. These processes are defined on an underlying probability space $(\Omega, \mathscr{F}, p)$. If $x \in C$ then it can be thought of as a curve in $L_2(\Omega; dp)$.

Received November 11, 1975; revised May 14, 1977.

AMS subject classification: Primary 60G35.

Key words and phrases: Multiplicity, sampling, denseness, mean square error, innovations, continuity, communication systems, reconstruction error.

* This work was supported by NSF Grant GK-32709.

ISSN 0047-259X

Consider also the class E of processes with finite energy (i.e., finite trace, $\int_T E[x(t)]^2 \, dt < \infty$). Clearly E and C overlap but also have disjoint portions. If $x \in E$, then it can be thought of as a member of $L_2(\Omega \times T; dp \times dt)$. Without loss of generality and in order to avoid technical complications it may be assumed that $E[x(t)] = 0$ and that T, the parameter set, is just a finite interval $[0, T]$. The results extend to the case of infinite T, but some of the proofs require modification.

The first aim of this paper is to establish that the class C is dense in E under the appropriate L_2 norm defined by

$$\| x \| = \left\{ \int_T E[x(t)]^2 \, dt \right\}^{1/2} = \left\{ \int_T R(t, t) \, dt \right\}^{1/2}$$

where $R(t, s) = E[x(t) \, x(s)]$ is the autocorrelation function of the process. Finally the processes involved may be considered to be purely nondeterministic. If a deterministic component is present most of the results and the proofs remain unchanged.

2. The Denseness in E of the Class of Processes with $M = 1$

It will be shown first that the random processes with unit multiplicity are dense in the class of mean-square continuous processes, and then that the latter is dense in E. The first step in the proof is constructive and demonstrates the denseness by considering the "sample-and-hold" technique. The second step consists of direct use of measure theory results [15]. The first step follows as a specialization of the following theorem.

Theorem 1. *Let $E[x(t)]^2 < \infty$, $\forall t$, and $x(t)$ be purely nondeterministic. Let the impulse response function $h(t, \tau)$ of a causal, linear system satisfy*

(i) $h(t, t) \neq 0, \forall t$,

(ii) *smoothness and square integrability conditions to ensure that the output $\hat{x}(t)$ of the system to a sequence of samples $x(t_i)$ belong to C.*

Then the multiplicity of $\hat{x}(t)$ is equal to one.

Proof. Assumption (ii), which is not restrictive for the purpose of this paper and which can be found in detail form in [12], is made to guarantee that the output $\hat{x}(t)$ which is given by

$$\hat{x}(t) = \sum_{0 \leqslant \tau_i \leqslant t} h(t, \tau_i) \, x(\tau_i), \qquad t \in [0, T] \tag{2.1}$$

and where $\{\tau_i\}$ is a linearly ordered set, has a multiplicity representation, and thus to bypass lengthy, unnecessary derivations and details.

The sequence $\{x(\tau_i)\}$ is linearly equivalent [2, 3] to a "white noise" sequence $\{n_i\}$, that is $E[n_i^2] = 1$, and

$$x(\tau_i) = \sum_{j \leqslant i} c_{ij} n_j \tag{2.2}$$

with

$$H(x, i) = H(n, i) \qquad \text{for every } i \tag{2.3}$$

where $H(\cdot, i)$ denotes the Hilbert space spanned by the random variables $(\cdot)_j$, $j \leqslant i$. Such a white noise can be formed by the Schmidt orthogonalization procedure, as done in [16, 17].

In the following corollary the previous theorem is specialized to establish that a "sample-and-hold" version of any process in C has unit multiplicity regardless of the original value of the multiplicity of such a process.

COROLLARY 1. *Let $x(t) \in C$. Suppose $x(t)$ is sampled periodically at a fixed period ϵ. Define $\hat{x}(t)$ as the "sample-and-hold" version of $x(t)$, that is*

$$\hat{x}(t) = x(n\epsilon), \qquad n\epsilon < t \leqslant (n+1)\epsilon \qquad n = 0, 1, 2, \ldots .$$

Then $\hat{x}(t)$ has multiplicity one.

Proof. Obvious with

$$\begin{aligned} h(t, \tau) &= 1, \qquad t - \epsilon < \tau \leqslant t, \\ &= 0, \qquad \text{elsewhere,} \end{aligned}$$

which clearly satisfies the assumptions of the theorem. ∎

Next the denseness of unit multiplicity processes in the class of mean-square continuous processes will be established in a straightforward fashion in view of the above results.

THEOREM 2. *The class of unit multiplicity processes is dense in the class of mean-square continuous processes.*

Proof. First it should be repeated that denseness is conceived under the norm in $L_2(\Omega \times T; dp \times dt)$. By the previous results it follows that if $x(t)$ is mean-square continuous, its sample-and-hold version $\hat{x}(t)$ has of course multiplicity one and, because of the uniform continuity in the closed interval $[0, T]$, given $\delta > 0$ it is possible to determine a sampling period ϵ small enough so that

$$E[x(t) - \hat{x}(t)]^2 < \delta^2/T$$

and hence

$$\| x - \hat{x} \| < \delta. \quad ∎$$

Let it be noted that only a "thin" subset of the unit multiplicity class o processes was used in the preceding constructive proof of the denseness. Thi: motivates the suspicion that this class may indeed be dense in broader classe: of processes. That this is the case will follow in the sequel.

Also note that the above proof is for finite T. For infinite T the proof exploits the square integrability of the process (i.e., the finite-energy property) in the obvious way by using the fact that given $\epsilon > 0$, there exists a finite T such that $\int_{t>T} E[x(t)]^2\, dt < \epsilon$ and by applying the same proof as above for the finite region T.

The proof of the denseness in E will now be completed by demonstrating the denseness of mean-square continuous processes in E.

THEOREM 3. *The class of mean-square continuous processes is dense in the space of finite-energy processes* $L_2(\Omega \times T; dp \times dt)$.

Proof. The proof will pretty much follow standard measure and integration theory techniques and will be for the infinite T case to exemplify the previous comment on the matter. Let $x \in E$; then $x(t)$ has finite power (i.e., $E[x(t)]^2 < \infty$) for almost all t (Lebesgue measure). Thus for the L_2 norm used here one may assume finite power everywhere. Now, since $x(t)$ is a measurable Hilbert space valued function of t, given $\epsilon > 0$ there exists [15] a continuous process $x_0(t)$, also in E, such that

$$m(S_1) < \epsilon, \qquad m = \text{Lebesgue measure}$$

where

$$S_1 = \{t \mid E[x(t) - x_0(t)]^2 > \epsilon\}.$$

The complement of S_1 consists of a part S_2 of finite measure M_2 and of a part S_3 which is the set where $E[x(t)]^2$ itself is very small. There $E[x_0(t)]^2$ is also comparably small and so is $E[x(t) - x_0(t)]^2$, to the extent that

$$\int_{S_3} E[x(t) - x_0(t)]^2\, dt < \epsilon.$$

This follows [15] from the finite trace assumption on $x(t)$. Again by [15] it follows that the integral over S_1 is also arbitrarily small by adjusting the ϵ's appropriately. On the remainder denoted by S_2, where $E[x(t) - x_0(t)]^2$ is less than ϵ, the integral is less than $\epsilon \cdot M_2$. Thus the norm $\| x - x_0 \|$ can be made arbitrarily small and the denseness in E is established. ∎

It is also possible to show similarly denseness of the class of sample function continuous processes in E under the same norm using the same procedures in reverse order. It is straightforward then to show [18] that the unit multiplicity processes are dense in the above class by just observing that given a process in E with continuous sample functions its sample-and-hold version, which, as shown, has multiplicity one, can be made to lie arbitrarily close to it in the E-norm. However, some labor is required due to the possibly nonuniform in Ω continuity of $x(t)$.

3. A Special Case

In the preceding section it was shown how the sample-and-hold version of an arbitrary finite trace random process can approximate that process by sampling adequately fast. The idea presented so far was really motivated by a consequence of the properties of the innovations representation [19–21] which are valid only for a subclass of E. The way that denseness of unit multiplicity processes is established in that subclass is interesting in that it offers an alternate constructive procedure, and in that it offers more insight into the nature of multiplicity and captures the essence of this denseness. The fact that this procedure is not applicable presently to the entire class E appears to be due to the lack of available wider versions of the innovations results rather than due to conceptual differences and limitations.

In particular consider $x \in E$ with the added requirement that it be mean-square absolutely continuous with its mean-square derivative x' belonging to E as well. The parameter set may be infinite. First the following lemma will be shown.

Lemma. *The multiplicities of x and x' as defined above are equal.*

Proof. Under the assumption of absolute continuity of x it is true [15] that

$$x(t) = \int_0^t x'(s)\, ds. \tag{3.1}$$

The assumption of purely nondeterministic processes implies that $x(0) = 0$. The lower limit of the integral in Eq. (3.1) might as well be $-\infty$. It is then evident that

$$H(x, t) \subset H(x', t), \qquad \forall t \tag{3.2}$$

where $H(\cdot, t)$ denotes the Hilbert space spanned by the variables $\cdot(s)$, $s \leqslant t$.

Furthermire, since $x'(t) = \text{l.i.m.}_{\epsilon\downarrow 0}(x(t) - x(t-\epsilon))/\epsilon$ it follows that

$$H(x', t) \subset H(x, t), \qquad \forall t. \tag{3.3}$$

Equations (3.2) and (3.3) imply that

$$H(x, t) = H(x', t), \qquad \forall t$$

and hence, that the multiplicities of the two processes are equal. ∎

Let $w(t)$ be a process with orthogonal increments in E, orthogonal to $x(t)$, and such that

$$E[w(t)]^2 = \sigma^2(t)$$

where $\sigma(t)$ is a continuous function satisfying

$$\int_0^T \sigma^2(t)\, dt < \epsilon \tag{3.4}$$

for a given $\epsilon > 0$, and T possibly infinite. In other words $w(t)$ is a "small" process with orthogonal increments. The process $\hat{x}(t)$

$$\hat{x}(t) = x(t) + w(t) = \int_0^t x'(s)\, ds + w(t)$$

has finite energy as well, and it obeys

$$\| x - \hat{x} \| = \| w \| = \int_0^T \sigma^2(t)\, dt < \epsilon.$$

However, regardless of the common value of the multiplicities of x and x' it is known [1] that $\hat{x}$ is equivalent to some process with orthogonal increments[1] and that, hence, it has multiplicity one. This result was originally known to be true under more stringent conditions and in particular for $w(t)$ a Wiener process (which would preclude the validity of Eq. (3.4) for infinite T) as well as Gaussian x and x'. The reason for that was that the equivalence of $\hat{x}$ to a Wiener process was derived in the sense of equivalent σ-fields. For the equivalence of the linear spans of the two processes, however, one only needs finite energy of x' as was shown by Rosanov [1].

Thus the following theorem was proved.

THEOREM 3. *The class of unit multiplicity processes is dense in the class of mean-square absolutely continuous processes in E.*

Obviously this theorem does not add anything to the previous theorems, except that it demonstrates the denseness via an alternate route amounting to showing that the perturbation of a process of arbitrary multiplicity by additive "white noise" of arbitrarily "small" magnitude reduced the value of the multiplicity to unity.

4. APPLICATION

In various communication problems, such as source encoding, quantizing, sampling, and in general processing random waveforms through digital equipment the observed or the signal process is usually sampled and then processed

[1] Equivalence in the sense of equal σ-fields or equal linear spans generated by the pasts of the two processes.

to yield a discrete-time version of a desired quantity (typically a functional of the data; for example, an estimate of the signal). The latter is often filtered to produce ideally the continuous-time version of that quantity. It is well known that such a procedure fails to achieve complete accuracy due to errors generated by sampling and quantizing. Most of these errors have been studied [22-25] and analyzed and have been attributed to identified sources such as jitter, aliasing, etc.

As Cramer has pointed out [8, 9] a discrete-time process has always unit multiplicity, while a continuous-time process, even if it is well behaved in terms of smoothness properties of its sample functions and its second-order statistics, may have arbitrary multiplicity. He has suggested that, therefore, care must be exercised in approximating continuous-time quantities with discrete-time data, since the multiplicity discrepancy may represent an additional source of error. It is evident from the discussion so far, that this should not be the case if, as it is commonly done, the mean-squared error is employed as a fidelity criterion. In this section it will be shown that, for a large class of acceptable processing transformations and for the case of integrated mean-square error as a figure of merit, the discrepancy in multiplicity values before and after digitizing does not contribute any additional amount of error. It should be pointed out that the suspicion for a separate multiplicity error is generated by the possibility that the reconstructed continuous-time version does not recapture the original value of the multiplicity, as indicated for example by Theorem 1. The case of the mean-square error alone as a criterion (rather than its integral) as well as the case of more general processing schemes will be subject of a future paper.

Let $x(t)$ be a random process that undergoes sampling, processing, and reconstruction. It will be assumed that $x \in E$. The output process $y(t)$ can be written as

$$y = F(x)$$

where F represents a general, possibly time-varying, nonlinear transformation with memory, like $y(t) = F(x(s); s \leqslant t, t)$. It will be assumed that F is continuous in the sense that

$$\| F(x_n) - F(x) \| \xrightarrow[n\to\infty]{} 0 \qquad \text{as} \quad \| x_n - x \| \xrightarrow[n\to\infty]{} 0$$

in the $L_2(\Omega \times T; dp \times dt)$ norm. Such a transformation can model accurately most sampling–estimation–reconstruction schemes used in practice. It may not be accurate for some elaborate source coding schemes of the predictive type [26, 27]. Of course it is further assumed that F maps into the same L_2.

Let the multiplicity M_x of the incoming process be strictly greater than one, while $M_y = 1$. (If instead it were only assumed that $M_x \neq M_y$ a triangular inequality type of approach would reduce the study of the problem to be present one.) As shown before there exists a process $\hat{x}$ with $M_{\hat{x}} = 1$ such that

$$\| x - \hat{x} \| < \epsilon.$$

Designate the output of F to $\hat{x}$ by $\hat{y}$. *Then*

$$y = F(x)$$

and

$$\hat{y} = F(\hat{x}).$$

By the assumed continuity of F, it is clear that

$$\| y - \hat{y} \| \to 0 \qquad \text{as} \quad \| x - \hat{x} \| \to 0.$$

However, it is true that

$$| \| y - x \| - \| \hat{y} - \hat{x} \| | \leqslant \| y - x - \hat{y} + \hat{x} \| \leqslant \| y - \hat{y} \| + \| x - \hat{x} \|.$$

Therefore as x is approximated by $\hat{x}$ that has unit multiplicity, the right-hand side of the above inequality goes to zero, and consequently the difference between the errors $\| y - x \|$ and $\| \hat{y} - \hat{x} \|$ is vanishing. Note that $\| y - x \|$ represents the integrated mean-square error for the actual state of affairs and therefore would include any amount of additional error due to multiplicity discrepancy. The error $\| \hat{y} - \hat{x} \|$ represents the hypothetical situation of processing $\hat{x}$ through F. If F is of the sampling and reconstruction type, then, as shown for example in Theorem 1, $M_{\hat{y}} = 1$ also. If not, then a new $\hat{y}$ can be found that has unit multiplicity and is essentially equal to the previous $\hat{y}$ as implied by the denseness properties established in the previous sections. Thus the error $\| \hat{y} - \hat{x} \|$ does not contain a component solely due to multiplicity discrepancy. Since the two errors can be made essentially equal it follows that the change in multiplicity introduced by F has a vanishing effect in the overall error.

5. Conclusion

The present understanding of the nature of the multiplicity M of a random process is rather incomplete. It is not clear what the consequences of the value of M are in terms of the behavior of the process. It is known of course that it involves only second-order properties, and that the autocovariance function is the only quantity determining the multiplicity value. In this paper it was shown that in terms of standard L_2-type of metric a random process can be thought of as having unit multiplicity. It still appears interesting to look at the $L_2(\Omega, dp)$ metric and show denseness with respect to that metric. At least for the special case treated earlier and concerning mean-square differentiable processes this appears to be true.

The application of the multiplicity representation to problems of engineering concern has been suspected all along and recently it has been demonstrated in a number of cases. In this paper it was shown that for suitably well-behaved modes of signal processing the act of digitizing and reconstructing that may affect the multiplicities of the waveforms, does not introduce significant additional error.

Finally the use of recent innovations results in establishing the denseness property of unit multiplicity processes has emphasized the relationship of the prewhitening of a process to the value of its multiplicity, and the mechanism that governs this relationship.

References

[1] Rosanov, Y. A. (1972). Innovation and non-anticipative processes. In *Multivariate Analysis*-III (P. R. Krishnaiah, Ed.), Academic Press, New York.

[2] Cramer, H. (1971). *Structural and Statistical Problems for a Class of Stochastic Processes*. Princeton Univ. Press, New Jersey.

[3] Cramer, H. (1961). On the structure of purely non-deterministic stochastic processes. *Ark. Mat.* **4** 249–266.

[4] Kallianpur, G. and Mandrekar. V. (1965). Multiplicity and representation theory of purely non-deterministic stochastic processes. *Theor. Probability Appl.* **10** 553–581.

[5] Ivkovich, L. and Rosanov, Y. A. (1971). On the canonical Hida-Cramer representation of stochastic processes. *Theor. Probability Appl.* **16** 348–353.

[6] Mandrekar, V. (1968). On multivariate wide-sense Markov processes. *Nagoya Math. J.* **33** 7–19.

[7] Ephremides, A. and Thomas, J. B. (1972). Multiplicity considerations in linear least-squares estimation. *SIAM J. Appl. Math.* **22** 162–172.

[8] Cramer, H. (1960). On some classes of nonstationary stochastic processes. In *Proceedings of the 4th Berkeley Symposium Statistics and Applied Probability*, Vol. 2, pp. 57–78.

[9] Cramer, H. (1964). Stochastic processes as curves in Hilbert space. *Theor. Probability Appl.* **9**, 169–179.

[10] Hida, T. (1960). Canonical representations of Gaussian processes and their applications. *Mem. Coll. Sci. Univ. Kyoto Ser. A* **33** 109–155.

[11] Lee, K. Y. and Mandrekar, V. (1973). Multiplicity and martingale approach to linear estimation. In *Proceedings of the 7th Annual Princeton Conference*, pp. 136–141.

[12] Ephremides, A. and Thomas, J. B. (1970). On the application of random signals with multiplicity one. In *Proceedings of the 8th Annual Allerton Conference*, pp. 589–598.

[13] Ephremides, A. and Brandenburg, L. (1973). On the reconstruction error of sampled data estimates. *IEEE Trans. Information Theory* **19** 365–387.

[14] Ephremides, A. and Thomas, J. B. (1973). *Random Processes—Multiplicity Theory and Canonical Decompositions*. Dowden, Hutchinson, and Ross, Stroudsburg.

[15] Royden, H. L. (1963). *Real Analysis*, Macmillan, New York.

[16] Ephremides, A. and Thomas, J. B. (1974). On random processes linearly equivalent to white noise. *Inform. Sci.* **7** 133–165.

[17] Levy, P. (1956). Sur une classe de coures de *l*'espace de Hilbert et sur une equation integrale non-lineaire. *Ann. Sci. Ecole. Norm. Sup.* **73** 121–156.

[18] Brandeburg, L. (1973). Private communication.

[19] Kailath, T. (1970). The innovations approach to detection and estimation. *Proc. IEEE* **58** 680–695.

[20] KAILATH, T. (1968). An innovations approach to least-squares estimation Part I: Linear filtering in additive white noise. *IEEE Trans. Automatic Control* **13** 646–655.
[21] SHEPP, L. A. (1966). Radon–Nikodym derivatives of Gaussian measures. *Ann. Math. Statist* **37** 321–354.
[22] KAHN, R. E. AND LIU, B. (1964). Sampling representation and optimum reconstruction of signals. *Communications Laboratory Technical Report* 7, Princeton University.
[23] SHAPIRO, H. S. AND SILVERMAN, R. A. (1960). Alias-free sampling of random noise. *SIAM J. Appl. Math.* **8** 225–248.
[24] LANDAU, H. J. (1967). Sampling, data transmission, and the Nyquist rate. *Proc. IEEE* **55** 1701–1706.
[25] BEUTLER, F. J. (1966). Error-free recovery of signals from irregularly spaced samples. *SIAM Rev.* **8** 328–335.
[26] BERGER, T. (1971). *Rate Distortion Theory*. Prentice–Hall, New York.
[27] TOMS, W. AND BERGER, T. (1971). Information rates of stochastically driven dynamic systems. *IEEE Trans. Information Theory* **17** 113–114.